Electronically Filed
Docket: 14-CRB-0011-SD (2010-2013)
Filing Date: 10/17/2019 07:49:27 PM EDT

Satellite Allocation Exhibit 1031

Before the
COPYRIGHT ROYALTY JUDGES
Washington, DC

In re	)	
DISTRIBUTION OF CABLE ROYALTY FUNDS	)	**NO. 14-CRB-0010-CD (2010-13)**

Written Rebuttal Testimony of

NANCY A. MATHIOWETZ
September 15, 2017

TABLE OF CONTENTS

Page

I. QUALIFICATIONS

1. I am Professor Emerita, Department of Sociology at the University of Wisconsin-Milwaukee (UWM). Prior to joining the faculty at UWM in 2003, I was Associate Professor, Joint Program in Survey Methodology, University of Maryland and University of Michigan. My research focuses on various aspects of survey methodology, including, but not limited to, the effects of mode and methods of data collection, question and questionnaire design, response error, and means to assess and reduce various sources of error in the survey process. I have taught courses on survey methodology, questionnaire design, and advanced statistical methods and have offered short courses on questionnaire design to various audiences. I have testified as an expert on survey research methodology in federal and state court cases.

2. My qualifications as an expert on survey research methodology are set forth in greater detail in Appendix A to my written direct testimony in this proceeding on behalf of the Joint Sports Claimants (JSC) (dated December 22, 2016).

II. INTRODUCTION AND SUMMARY

3. My written direct testimony discusses the 2010-13 cable operator surveys conducted by Bortz Media & Sports Group, Inc. (Bortz surveys). As I explain in that testimony, the Bortz surveys provide a valid and reliable assessment of the relative market value of the different categories of distant signal programming that cable systems carried during the years 2010-13. The purpose of my rebuttal testimony is to address the written direct testimony submitted in this proceeding by (1) Joel Steckel, Ph.D., Howard Horowitz, and Martin R. Frankel, Ph.D., on behalf of the Program Suppliers; and (2) Debra J. Ringold, Ph.D. on behalf of the Canadian Claimants Group.

4. The testimony of Dr. Joel Steckel is a critique of surveys of cable system executives, that is, the methodology used by both Horowitz and Bortz in their respective data collection efforts. In my opinion, Dr. Steckel is incorrect to assert that cable operator surveys are inadequate for assessing the issue of relative market value in this proceeding. Dr. Steckel's criticism are far ranging; he asserts that the surveys do not measure market value, sample the wrong population (cable system operators), and result in invalid data due to the nature of the key constant sum question (which he considers too complex) and the mode of data collection (telephone). These criticisms have been raised in previous proceedings; the Copyright Royalty Judges (CRJs) in the Distribution of the 2004 and 2005 Cable Royalty Funds noted, "Yet, whether taken individually or viewed as a group, we do not find these other criticisms to undermine the general usefulness of the Bortz survey for the purpose offered" (Federal Register, Vol. 75, September 17, 2010, p. 57068). I provide below (see Section III) detailed responses to Dr. Steckel's arguments against the use of the Bortz survey data.

5. The testimony of Howard Horowitz and Dr. Martin Frankel present the methodology and findings from surveys conducted in 2010-2013 ("Horowitz surveys"); the methodology used in the Horowitz surveys is similar to that used by Bortz for the JSC. However, there are key differences in the design and implementation of the Horowitz survey and the Bortz survey that I discuss below. The testimony of Dr. Debra Ringold describes the methodology and findings from surveys conducted in 2010-2013; in contrast to the Bortz and Horowitz surveys, the Ringold/Ford surveys are limited to the assessment of the relative value of programming on Canadian Signals.

6. While properly designed cable operator surveys are useful for assessing relative value in this proceeding, my review of the Horowitz survey and the Ford/Ringold survey leads me to conclude that the flaws in each of these surveys renders them neither reliable nor valid for the production of valuation estimates. As detailed below (Section IV), the Horowitz survey design suffers from a number of significant flaws, most notably the inclusion of incorrect and misleading information as part of the questions posed to the respondents. In addition, the implementation methodology places undue burden on the respondents, asking executives to provide information for the full universe of CSOs (not just the sampled CSOs) as well as asking executives to report about a large number of CSOs, often in a single questionnaire.

7. With respect to the Ford/Ringold survey, the analytic sample is biased, giving preference to French-language systems, and its small sample size leads to unreliable estimates. Other concerns with the Ford/Ringold survey are detailed below (Section V).

III. DR. STECKEL'S CRITICISMS OF THE BORTZ SURVEY ARE WITHOUT MERIT

8. Dr. Joel Steckel criticizes both the Horowitz and Bortz surveys. He asserts that the surveys do not measure market value, sample the wrong population (cable system operators), and result in invalid data due to the nature of the key constant sum question (which he considers too complex) and the mode of data collection (telephone). Dr. Steckel advocates for surveying the consumers of cable system programming, the subscriber, as opposed to surveying cable system operators. These are not new arguments in these proceedings—for example, each of these points was previously made

by Program Suppliers' expert Dr. Alan Rubin, whom Dr. Steckel cites (p. 34)[1]—and despite these arguments the CRJs, their predecessors and the courts repeatedly have found the Bortz survey to be useful in determining the appropriate allocation of copyright royalties.

9. I disagree with Dr. Steckel's assessment of the two surveys.[2] In reviewing Dr. Steckel's critique, I will draw upon Diamond's "Reference Guide on Survey Research," one of the chapters of the *Reference Manual on Scientific Evidence, Third Edition* (2011).[3] Diamond frames her chapter as responses to a series of questions, several of which speak directly to the concerns raised by Dr. Steckel. These questions include:

- Was the survey designed to address relevant questions?
- Was an appropriate universe or population identified?
- Were questions on the survey framed to be clear, precise and unbiased?
- What limitations are associated with the mode of data collection used in the survey?

Dr. Steckel also raises other concerns that do not align with the *Reference Manual.* I will address these issues at the end of this section.

---

[1] Like Dr. Steckel, Dr. Rubin argued that the appropriate population to survey was not cable system operators but cable subscribers (*e.g.*, September 2009 Corrected Testimony of Alan M. Rubin, pp. 4, 9-14); that the Bortz constant sum question was too complex (*e.g.*, November 1991 Testimony of Alan M. Rubin, pp. 10-11; October 1985 Testimony of Alan M. Rubin, pp. 5-6); and that the surveys should not have been conducted over the telephone (*e.g.*, November 1991 Testimony of Alan M. Rubin, p. 7).

[2] I note that Dr. Steckel's review of the Bortz survey relies on the 2004-2005 Bortz surveys and does not reflect multiple changes made in the methodology for the 2010-2013 Bortz surveys, and therefore a number of his criticisms are inapplicable to the Bortz surveys at issue in these proceedings.

[3] Dr. Steckel cites to a brief discussion of survey research in the *Manual for Complex Litigation* (4th ed. 2004), which includes some similar criteria to, but is less comprehensive than, Diamond's chapter in the 2011 *Reference Manual.*

A. Was the survey designed to address relevant questions?

10. The language used by the CRJs in the Distribution of the 2004 and 2005 Cable Royalty Funds (Federal Register, Vol. 75, September 17, 2010) states:

> …the sole governing standard is the relative marketplace value of the distant broadcast signal programming retransmitted by cable systems during 2004 and 2005 (p. 57065).

Dr. Steckel asserts that the Bortz and Horowitz surveys' measurements of the cable system operators' valuations do not correspond to the marketplace value standard. As Dr. Steckel acknowledges (p. 22), the Copyright Arbitration Royalty Panel (CARP) determined that the constant sum question posed in the Bortz survey "is largely the question the Panel poses when it constructs a simulated market" (Report of the CARP in Docket No. 94-3 CARP CD 90-92, p. 65 (May 31, 1996).[4] The CARP further stated that the Bortz survey was "focused more directly than any other evidence to the issue presented: relative market value" (*Id.*).

11. Dr. Steckel contends that the CARP was incorrect. However, subsequent decisions in statutory royalty proceedings likewise have found that the Bortz survey is well-suited to assessing the relative market value of different types of programming to cable system operators (CSOs) in the hypothetical market. For example, in approving the CARP allocation of the 1998-99 cable royalties, the Librarian of Congress approved the CARP's reliance on the Bortz survey and cited the CARP's determination "that the Bortz survey best projected the value of broadcast programming in the hypothetical

[4] The Librarian of Congress adjusted the CARP's royalty allocations to account for settlements of claims by the Music Claimants and National Public Radio, and to correct errors in the apportionment of "3.75 Funds," and otherwise approved the CARP's determination; the Librarian's decision was affirmed on appeal. *National Association of Broadcasters v. Librarian of Congress*, 146 F.3d 907 (D.C. Cir. 1998).

marketplace . . ." (Federal Register, Vol. 69, January 26, 2004, p. 3609). The Librarian's decision was affirmed in an appellate decision stating: "Nor did the CARP act unreasonably in declining to rely on Nielsen for direct evidence of viewing, as Bortz adequately measured the key criterion of relative market value. Moreover, as the CARP put it, Bortz 'subsumes *inter alia* all viewing data that a CSO might consider when assessing relative value of programming groups.'" *Program Suppliers v. Librarian of Congress*, 409 F.3d 395, 402 (D.C. Cir. 2005). The court further observed that "[t]he Bortz survey, supplied by JSC, measures what CSOs perceive as the relative market value of different types of programming." *Id.* at 398. Similarly, the CRJs' decision allocating the 2004-05 cable royalties found "the Bortz study to be the most persuasive piece of evidence provided on relative value in this proceeding" and that "[t]he Bortz intervals certainly mark the most strongly anchored range of relative programming values produced by the evidence in this proceeding" (Federal Register, Vol. 75, September 17, 2010, pp. 57066, 57068).

12. Based on the historical comments of the CRJs, CARP, the Librarian, and the Court of Appeals, it appears that both the Bortz and Horowitz surveys, by focusing on the relative valuations placed on program categories by cable system operators, are in fact addressing the relevant question of interest.

B. Was an appropriate universe or population identified?

13. Dr. Steckel criticizes both the Bortz and Horowitz surveys for surveying cable system operator executives. Specifically, he maintains that "the relevant opinions for projecting marketplace results are not those of cable executives; they are those of cable customers" (p. 40). He goes on to state, "If you want to know what customers (*i.e.*, the market) value, ask them" (p. 41). However, as discussed above, the CRJs, CARP, the

Librarian and the appellate court consistently have stated that the relevant customers in the hypothetical market would be the CSOs, and that the Bortz survey is an appropriate methodology for assessing CSOs' relative valuations. Thus, the CRJs' 2004-05 determination stated "the Bortz survey focuses on the appropriate buyer in the hypothetical market—*i.e.*, the cable operator" (Federal Register, Vol. 75, September 17, 2010, p. 57066).

C. Were questions on the survey framed to be clear, precise and unbiased?

14. The criticisms that Dr. Steckel offers with respect to the constant sum questions are unfounded. As the Librarian has observed, "'uncontroverted testimony and years of research indicate rather conclusively that constant sum methodology, as utilized in the Bortz survey, is highly predictive of actual marketplace behavior'" (Federal Register, Vol. 69, January 26, 2004, p. 3615). The CRJs have likewise stated: "We find that the Bortz study is founded on a method—the constant sum survey—that has been long regarded as a recognized approach to market research. Tr. at 50 (Trautman), 1299 (Ringold), and 3007 (Gary Ford)" (Federal Register, Vol. 75, September 17, 2010, pp. 57066-67). These findings reflected substantial evidence presented by JSC and other parties regarding the suitability of constant sum questions for purposes of the Bortz survey.

15. For example, as Dr. Steckel notes (p. 34), Professor Leonard Reid presented detailed testimony explaining why constant sum questions were appropriate for the Bortz survey. Professor Reid explained that "[t]he constant sum scale is a widely accepted and often-used measurement tool in marketing research" and discussed a number of the underlying studies establishing the utility of that technique (August 1991 Testimony of

Leonard N. Reid (Reid Testimony) (JSC Written Direct Statement, Vol. II, Tab 14), pp. 5-14). As Professor Reid observed, "the constant sum technique is particularly well-suited for measuring behavioral intentions, past actions, and evaluative preferences" (Reid Testimony, p. 6). He further observed that "[t]he pragmatic value of the constant sum technique for measurement purposes may be demonstrated by its application in the field," noting the routine use of this technique by leading marketing firms and major corporations (Reid Testimony, pp. 12-14).

16. While Dr. Steckel faults Dr. Reid for citing (among other studies) a seminal study by Dr. Joel Axelrod and suggests that Dr. Axelrod's study weighs against the use of the constant sum technique for purposes of the Bortz survey (p. 35), he ignores (and perhaps was unaware) that Dr. Axelrod himself has testified in a prior cable royalty distribution proceeding that "the use of the constant sum technique in order to determine the relative values that cable operators attach to different types of programming is appropriate" and that nothing in his study suggests any issue with Bortz's use of that technique (Oral Testimony of Joel N. Axelrod, Docket No. 94-3 CARP CD 1990-1992 (Axelrod Oral Testimony) (JSC Written Direct Statement, Vol. III, Tab 2), pp. 11130-34, 11249-50; February 1996 Rebuttal Testimony of Joel Axelrod (Axelrod Rebuttal Testimony) (JSC Written Direct Statement, Vol. II, Tab 2), pp. 2-4).

17. I agree with Dr. Steckel that the constant sum question might be difficult to answer if posed to respondents of a general population survey. But the respondents to the Bortz and Horowitz surveys are executives of cable system operations, who engage in complex business decisions as part of their professional lives. Dr. Steckel suggests that the task in the constant sum method requires executives to make judgments about

"unfamiliar constructs," but program valuations are not unfamiliar constructs to cable system executives.[5] As noted by Bortz, survey interviewers sought responses from the individual "most responsible for programming carriage decisions" (Bortz, pp. 14-15). The Written Rebuttal Testimony of Daniel Hartman (pp. 16-17) and Allan Singer (p. 11) confirm that the task of assessing relative value of programs is part of the job related to purchasing signals.

18. Dr. Steckel also fails to account for differences between the Bortz and Horowitz surveys with respect to the formulation of the questions. It is important to point out that in his critique of the Bortz methodology, Dr. Steckel reviewed the 2004-2005 data collection instrument and not the revised instrument used by Bortz for the 2010-2013 surveys. Presented below is the wording of the constant sum question used by Bortz in 2010-2013:

> Now, I would like you to estimate the relative value to your cable system of each category of programming actually broadcast by the stations I mentioned during 2010, excluding any national network programming from ABC, CBS and NBC. Just as a reminder, we are only interested in U.S. commercial station(s) _____, U.S. non-commercial station(s) _____, and Canadian station(s) _____.... Assume your system spent a fixed dollar amount in 2010 to acquire all the non-network programming actually broadcast during 2010 by the stations I listed. What percentage, if any, of the fixed dollar amount would your system have spent for each category of programming? Please write down your estimates, and make sure they add to 100 percent.[6]

[5] I note that this argument has been asserted previously. See October 1985 Testimony of Alan M. Rubin (p. 5) in which he states, "Operators and subscribers were asked to do something completely abnormal to their routine cable television behaviors." Despite this criticism, previous CRJs have consistently looked to the Bortz survey with respect to their allocation decisions (see, for example, Federal Register, Vol. 75, September 17, 2010).

[6] In response to comments expressed by the CRJs in their 2004-2005 Distribution Order, the wording used in 2010-2013 was modified from the wording used in 2004 and 2005 where, as in previous surveys, the Bortz constant sum question asked respondents to "assess the different programming categories in terms of their relative value in 'attracting and retaining subscribers'" (Bortz, p. 40).

The wording used for the Horowitz surveys is as follows[7]:

> Now, considering everything we have been discussing, I would like you to estimate the ***relative value*** to your cable system of each type of **[NETWORK CARRYING SYSTEMS (E): *non-network*]** programming actually broadcast during 2013 by [INSERT STATION(S) FROM LIST - COLUMN J]…. Assume you had a fixed dollar amount to allocate for the **[NETWORK CARRYING SYSTEMS (E): *non-network*]** programming actually broadcast during 2013 on [INSERT STATION(S) FROM LIST - COLUMN J]…. Considering the value of each type of programming to your cable system, what percentage, if any, of the fixed dollar amount would you allocate for each type of programming? Please write down your estimates and make sure they add to 100 percent.… In formulating your percentage, please think about all the factors we have been discussing, including using this programming in your advertising and promotions in 2013 to attract and retain customers, the importance of this programming to you and your subscribers, and any other considerations you may have.

As is evident from a comparison of the wording of these two constant sum questions, the Horowitz question asks the respondent to focus on valuations related to advertising and attracting and retaining customers, similar to the wording used in 2004-2005 by Bortz and criticized by the CRJs with respect to the 2004-2005 Distribution of Cable Royalty Funds. While the Horowitz question used in 2010-2013 does ask the respondent to "think about all factors," the wording specifically calls out the issue of attracting and retaining customers. As noted by the CRJs in 2010, "a myriad of other net revenue considerations may be involved in any programming decisions" (Federal Register, Vol. 75, September 17, 2010, p. 57066).

19. A key requirement as outlined by Diamond is that questions be framed so as not to bias the respondents. As discussed in part IV below, the Horowitz questionnaire fails this condition, specifically in its use of examples for the Program Suppliers category.

[7] Note that the wording reported here is not the wording used for PBS only or Canadian only stations. See Direct Testimony of Howard Horowitz (Horowitz), Appendix A, p. 36.

The examples used to describe the Program Suppliers category are misleading and biased in favor of Program Suppliers. See pages 16-18 below.

20. Dr. Steckel states that both the Bortz and Horowitz questions are "ambiguous" (p. 25) because the respondent is asked about a "system" (singular) when, in many cases, the respondent has responsibility for multiple cable systems. However, on this design issue, the Bortz and Horowitz surveys differ significantly. In the Bortz survey, if a single executive was responsible for more than one cable system, that executive completed a separate survey questionnaire for each system, focusing on a single cable system's distant signals for each questionnaire (Written Rebuttal Testimony of James M. Trautman, p. 43, n. 29). In contrast, in the Horowitz survey, when a single executive was the respondent for more than one system, the executive "was only asked to respond to one survey for all the systems with the same channels" (Horowitz, p. 8), meaning that the respondent was tasked with addressing multiple cable systems in a single survey questionnaire. Hence, the criticism offered by Dr. Steckel on this point is only applicable to the Horowitz data collection effort.

21. I note that Dr. Steckel offers no empirical data to support his assertion that the constant sum questions are "complex" (p. 28).[8] In my experience, when respondents are asked questions that they are not able to process cognitively due to the complexity of the question, the data reflect this in either high rates of missing data or illogical responses. We see neither of these patterns in the Bortz data.

[8] Program Suppliers' experts have made the same assertion in prior proceedings; see for example the October 1985 Written Direct Testimony of Alan M. Rubin, and November 1991 Written Direct Testimony of Alan M. Rubin. Despite these previous concerns, the Program Suppliers adopted a constant sum methodology for the measurement of valuation in 2010-2013.

22. Finally, in his criticism of the constant sum methodology, Dr. Steckel notes several recent publications that outline new methodologies for collecting preference data. In contrast to the vast literature supporting the extensively used constant sum approach, Dr. Steckel is advocating for the adoption of techniques only recently introduced in the literature without significant testing and validation for the question of interest to the CRJs.

23. With respect to the Lourviere and Islam article cited by Dr. Steckel for the proposition that "indirect" measures of importance outperform direct measures, it is important to note that the authors also offer several cautions with respect to the use of "indirect" measures of which Dr. Steckel is advocating, including the susceptibility of these measures to context effects. Moreover, the authors never conclude that indirect measures outperform direct measures such as constant sum questions.

24. With respect to the other methodologies cited by Dr. Steckel (Netzer and Srinivasan, 2011 and Srinivasan and Wyner, 2009), these studies have only recently moved into the peer-reviewed literature, and both studies are based on web-based data collection (no interviewer) and focus on cases where there are a large number of attributes to assess (> 10). In contrast, the Bortz and Horowitz constant sum task focuses on only 5 to 8 program categories (depending upon the system) and were completed through live telephone interviews. One would be remiss to adopt the new approaches described in these articles based on the findings from a few recent studies.

D. What limitations are associated with the mode of data collection used in the survey?

25. Dr. Steckel claims that using the telephone for data collection results in unreliable and invalid data. Yet he does not provide any empirical support for that claim, and he

ignores that telephone surveys of business entities are widely used and recognized as producing reliable, valid data.[9]

26. As Dr. Axelrod testified in the 1990-1992 royalty proceedings, the use of telephone surveys is "an accepted survey research technique," is "widely done," and is appropriate for the purpose of administering the Bortz survey (see Axelrod Oral Testimony, pp. 11122, 11130-11134, 11223-25). The decision as to which mode of data collection to use is one that concerns tradeoffs between costs and potential errors. Each mode has its benefits and its limitations. Self-administered surveys such as those conducted via traditional mail or as web-based surveys benefit from allowing the respondent to read the material but are limited in that (1) one is never assured that the respondent fully reads any one question; (2) one cannot know with certainty who has served as the respondent; and (3) the lack of an interviewer forces the respondent to undertake the task by him/her self, with no means to seek clarification concerning a question or a response category. Interviewer-administered questions benefit from the presence of an interviewer—both to encourage overall response and to assist in the task—but the presence of an interviewer can also be detrimental in the measurement of socially desirable or undesirable behavior.

27. Indeed, the use of the telephone for the collection of survey data has been popular in the United States since the early 1970s and only recently has been in decline for general population surveys. However, for the Bortz and Horowitz surveys, we are not discussing general population surveys but rather a survey of business entities for which

[9] I note that Alan Rubin in his Testimony of November, 1991 also asserted that the constant sum technique should only be used with "personal, face-to-face interviewing" (p. 7).

telephone surveys are quite prevalent. Moreover, to assist respondents at CSOs who carried only WGNA distant signals, the Bortz methodology used for the 2010-2013 data collection included advance mailing of materials identifying the compensable and non-compensable programming on WGNA. In contrast, the Horowitz survey did not provide such materials. As a result, no clear delineation of compensable and non-compensable programs was articulated for respondents to the Horowitz survey for whom WGNA was the only distant signal carried.

28. Dr. Steckel also criticizes the use of telephone surveys for data collection, citing a paper by Dr. Joel Axelrod as "caution[ing] against using constant sum measures in a telephone interview" (p. 35). However, in prior proceedings Dr. Axelrod himself appeared as a witness, discussed that same paper, and testified that the use of telephone surveys was appropriate for the purpose of administering the constant sum question in the Bortz survey (see Axelrod Oral Testimony, pp. 11130-11134).

29. I note that Dr. Steckel incorrectly asserts that the unit of analysis of the Bortz and Horowitz surveys is the cable system executive and not the cable system. He states: "The data are collected and tabulated with the unit of analysis being the respondent cable system executive, not the cable system" (p. 25). While the *respondent* in each of the surveys is an executive, the *analytic unit* for each of the surveys is the cable system, with weights corresponding to copyright royalties paid by the system. Based on his comments, it appears that Dr. Steckel has not examined the data from either the Bortz or Horowitz data collection efforts. Dr. Steckel is incorrect in his assertion that estimates from the studies are biased in favor of small cable operators.

30. In sum, I find the arguments put forth by Dr. Steckel to reiterate previous concerns expressed by experts for the Program Suppliers and which, in previous proceedings, have not been found to undermine either the methodology of or the estimates derived from the Bortz survey. I disagree with Dr. Steckel's assessment that the Bortz and Horowitz surveys focus on the wrong population to study; he asserts that the viewing public and not cable system executives should be the focus of study. Cable system executives *are* the relevant population to study for this task; in contrast to the viewing public, CSO executives are familiar with the concept of program valuations and utilize this information in contract negotiations. As such, there is no foundational support for Dr. Steckel's criticism that the constant sum question is "too complex."

IV. THE HOROWITZ SURVEY IS FUNDAMENTALLY FLAWED AND PROVIDES NEITHER A VALID NOR RELIABLE BASIS FOR ESTIMATING RELATIVE VALUE

31. The written direct testimony of Howard Horowitz summarizes the design and implementation of cable system operator surveys conducted by Horowitz Research for each of the years 2010-2013. The written direct testimony of Martin R. Frankel, Ph.D. provides information related to the sample design and estimation for the Horowitz surveys, 2010-2013.

32. The questionnaire and sample design of the Horowitz survey are similar in nature to those used by Bortz Media and Sports Group, Inc. Both surveys make use of a stratified random sample of Form 3 cable system operators, for which the strata are defined according to annual royalty amounts for the respective years. The mode of data collection is the same for the two studies—telephone—and the key question of interest, that is, program valuation, is based on a constant sum methodology. The survey questionnaire for both Bortz and Horowitz includes preliminary questions that measure

the respondent's perception of the importance of different types of program categories and introduces the respondent to the specific program categories of interest. The implementation of the two studies calls for both interviewers and respondents to be blinded to the respective sponsors of the data collection effort. And in the implementation of the two sets of studies, we see response rates that exceed the current norms in the industry.

33. However, there are significant differences in the two studies, and these differences are critical to understand in assessing the relative validity and reliability of the two sets of estimates for 2010-2013. The key design differences between the Bortz and Horowitz surveys include the following:

- The misuse of illustrative programming examples and "such as" programming descriptions—including the provision of incorrect examples, incorrect descriptions and programs that were not broadcast on a compensable basis;
- The failure to provide information identifying compensable programs on WGNA;
- The addition of an inappropriate "other sports programming" category;
- The consolidation of surveys in which a respondent was queried about multiple systems simultaneously; and
- The unnecessary burden of requiring respondents to consider *all* of the distant signals carried by a cable system.

A. Misuse of Illustrative Examples and "Such As" Descriptions

34. The Horowitz survey's relative value question (Question 6) violates general principles of questionnaire design due to its misleading and inconsistent use of examples and "such as" descriptions across program categories. As discussed in Diamond's "Reference Guide on Survey Research," a fundamental requirement for a sound survey is that the questions must be "clear, precise and unbiased" (p. 387). Even an accurate example may inject bias into a survey question—for example by limiting respondents'

consideration to those examples that are offered (Beatty, Cosenza, and Fowler, 2006), or by increasing the reported frequency for the response category (Tourangeau, Conrad, Couper, and Ye, 2014). And where a survey question uses an inaccurate or misleading example, that renders the question (of which the example is part) inherently imprecise and biased. If examples are meant to serve as a means to improve comprehension of a question or a response category, then it is imperative that the examples not be misleading.

35. Of the problems with the Horowitz survey's relative value question, the inclusion of incorrect information as part of the response category descriptions is the most egregious. The rebuttal testimony of James Trautman lists in detail numerous errors in the program examples and "such as" program descriptions provided to the Horowitz survey respondents, both with respect to all of the WGNA-only systems and systems that included only WGNA and public broadcasting, as well as many of the other systems (Written Rebuttal Testimony of James M. Trautman, pp. 18-28). These errors include providing the cable system respondents with examples and descriptions of programming that the cable systems did not actually carry, or that was not compensable, or that was attributed to the incorrect program category. As a result of these inaccuracies, the questions are biased and therefore the responses are not valid representations of valuations for the various program categories.

36. In addition to these errors, I also note that the descriptions of program categories are inconsistent across the categories. As shown in Appendix A of Horowitz, no examples are offered with respect to the category "News and Community Events," whereas a similarly self-explanatory category "Movies" offers six examples in addition to three movie sub-categories offered as part of the "such as" clause. The examples offered

for the “Live, play-by-play coverage of professional and college team sports” are not examples but rather the full enumeration of the sports programs associated with this category. Sometimes a program category includes examples of sub-categories (through the use of “such as” descriptions) as well as specific program titles; for other program categories there are neither examples of sub-categories nor examples of specific program titles; and some program categories include only specific program examples.

37. These inconsistencies in the program category descriptions are significant. First, respondents give greater cognitive processing the longer the response category offered—so those categories that incorporate “such as” program subcategories and illustrative examples will benefit from greater cognitive processing by the respondent. The goal in designing response categories for a question is to minimize differences in the level of cognitive processing by the respondent across the various categories since differences in the level of processing may contribute to differences in responses. Second, frequency—or in this case, relative valuations—most likely are impacted by the use of examples. Thus, we would expect that valuations across categories could have differed, in part, as a result of the variation in language (“such as”) and variation in the use of illustrative examples. So as to minimize the measurement error attributed to question wording, each of the program categories should have been treated equally with respect to the number of illustrative examples and the use of “such as” language.

38. Although the inconsistencies in the structure of the program categories most likely impacts the estimation for these respective categories, it is the presentation of misleading information included in the description of program categories that results in my assessment that the questions (and response categories) are biased.

B. Failure to Identify Compensable WGNA Programming

39. Not only is the valuation question flawed due to what information is provided, the Horowitz questionnaire also suffers from errors of omission, specifically with respect to the identification of compensable programs for systems that carried WGNA. A key issue for signals that carry WGNA is for the respondent to understand which programs on WGNA are compensable and which are not. The Bortz surveys of WGNA-only systems addressed this issue by pre-mailing affected respondents a description of the compensable programs on WGNA every year, including the total number of hours of such programming (see Bortz, p. 30).

40. This feature of the Bortz surveys was new to the 2010-2013 data collection effort and addresses, in part, a concern raised by the CRJs as part of the distribution of the 2004-2005 cable royalty funds (Federal Register, Vol. 75, September 17, 2010, p. 57067).[10] In contrast, the Horowitz survey merely instructed respondents with WGNA systems as follows: "Please do not assign any value to programs that are substituted for WGN's blacked out programming" (Horowitz, Appendix A, p. 36). Cable system operators, however, have no reason to know which programs on WGNA are or are not substituted for blacked-out programming of the local WGN-Chicago station (see Written Rebuttal Testimony of James M. Trautman, pp. 14-15; Written Rebuttal Testimony of Allan Singer, p. 8).

41. Of particular importance is the fact that all of the non-compensable programming on WGNA falls within the Program Suppliers and Devotional categories (Written

[10] I note that the list of compensable programs and hours of airtime were only sent to those systems for which WGNA was their only distant signal. Systems for which WGNA was one of two or more distant signals did not receive this information.

Rebuttal Testimony of James M. Trautman, p. 14). To the extent that the respondent does not fully understand and differentiate between compensable and non-compensable programs, the relative valuations for the Program Suppliers categories (movies, syndicated series, and "other" sports) as well as the Devotional category will be upwardly biased. Hence, I find that the methodology used by Bortz for WGNA-only (in which compensable programs were clearly delineated for the respondent) would lead me to conclude that for WGNA-only stations, the Bortz estimates would provide a more valid estimate of relative program valuations.[11]

C. Addition of "Other Sports Programming" Category

42. Another key difference between the Bortz and Horowitz surveys is the inclusion of an "Other sports" program category in the Horowitz survey. Treating a category as minor as "other sports" in the same manner as program categories such as "movies" and "live professional and college sports" suggests to the respondent that the category is significant and on par with the other seven categories. I agree with Mr. Trautman's assessment that the provision of these misleading examples would lead to inflated estimates of the relative value of "other sports." For example, if we look at those systems that retransmitted WGNA as their only commercial distant signal during 2010-2013, we see responses in the Horowitz data that are illogical, given the fact that WGNA carried less than two hours each year of compensable "Other Sports" (Trautman Written Rebuttal Testimony, p. 17). For example, in 2013, one of the responding CSOs assigned relative

[11] I note that for those cable systems for which WGNA is one of two or more distant signals carried, neither Bortz nor Horowitz provided respondents with a list of compensable programs. For those "WGNA-plus" systems, the Bortz surveys provide a more valid estimate of relative program valuations than the Horowitz surveys due to the flaws in the Horowitz WGNA-plus surveys discussed herein, such as the use of misleading and inaccurate program examples and the inappropriate addition of an "Other Sports" category.

valuations of '25' for both Live Team Sports and "Other Sports." Other examples include three responding CSOs that each valued Live Team Sports at '40' and "Other Sports" at '30' despite the fact that the only compensable "Other Sport" broadcast was a single one-hour horse race ("The Arlington Million") (Trautman Written Rebuttal Testimony, p. 17).

D. Respondent Selection

43. The Bortz and Horowitz data collection methodologies differed in their approach to identifying the respondent of interest and how interviews were conducted. For the Bortz study, interviewers sought to interview the individual "most responsible for programming carriage decisions" (Bortz, pp. 14-15). As noted by Bortz, "In attempting to reach this individual, the interviewer was frequently referred to a regional executive" (p. 15). As such, Bortz often began at the CSO level to identify the person responsible for programming and moved up to a regional executive when necessary. The Bortz approach of starting at the CSO level limited the number of cable systems for which a single executive served as a respondent to a maximum of eleven, with the average number of CSOs for which a respondent reported ranging between 2 (2011) and 2.4 (2010) and the modal number of responses being 1 (that is, most respondents only responded for one system) (Trautman Written Rebuttal Testimony, Table A-4). Moreover, when the same individual was selected to report on multiple cable systems, he or she was administered a separate questionnaire for each system so as to focus solely on a single cable system at a time.

44. The Horowitz survey methodology also calls for the selection of "the executive with the decision-making authority over the carriage of distant broadcast signals for each CSO in our sample" (Horowitz, p. 5). However, in contrast to the approach used by

Bortz, the methodology used by Horowitz begins at the top of the decision making process, often at the MSO level. As a result, some respondents had significant response burden, being asked to report on an extremely large number of cable systems. For example, we see that in 2013 the AT&T MSO includes 60 CSOs in the universe of systems surveyed by Horowitz, and that a single executive was interviewed with respect to all 60 CSOs (Horowitz, Appendix B, p. 41). Focusing on the Horowitz sample systems, the number of cable systems for which a single executive provided data was as high as 38 (in 2013).[12] Also in contrast to the Bortz methodology, in the Horowitz survey, when a single executive was responsible for multiple systems and each of those systems had the same distant channel lineup, then only a single survey was administered. (Horowitz, p. 8).

45. For these reasons, the Horowitz methodology places excessive burden on the selected respondent. For the Horowitz survey, an executive was asked to report not only about those cable systems selected for the sample, but also for all systems for which he or she was responsible in the *entire universe* of Form 3 cable systems that transmitted a distant signal (Horowitz, p. 8). As a result, you see the extremely high number of cable systems for which an individual had to respond evident in the tables of Appendix B of the Horowitz report. Rather than focus on those CSOs that form the basis for the estimation, a respondent had to evaluate a much larger set of CSOs to determine his or her program relative valuations. The task as posed in the Horowitz survey (asking a single individual respondent about many CSOs either in a single interview or across multiple interviews

[12] JSC_2010_2013_Masked_withDistantStations_MSOchanges_13July2017.xlsx.

for those cases with different distant signals) would lead respondents to make summary judgments concerning valuations.

46. These summary judgments, in the case of multiple CSOs with the same distant signal, will reflect valuations for *sampled* CSOs as well as *non-sampled* CSOs since Horowitz asked respondents to report on the universe of all CSOs.

47. The pooling of multiple CSOs with the same distant signal lineup into a single questionnaire assumes that the valuation for those distant signals is the same, regardless of the population being served by those distant signals. Consider, for example, the case of WGNA, a distant signal that is transmitted throughout the country. One can imagine that interest in the Chicago sports teams or Chicago-related news may be greater in some parts of the country than others. To group all of the WGNA systems together in requesting relative program valuations makes an assumption about the equality of valuations for every cable system that offers WGNA (among those reported for by the same respondent). Addressing multiple systems in a single survey meant the respondents had to somehow provide a single valuation for signals carried across a large number of systems that were likely geographically diverse.

48. In addition to the burden related to reporting for multiple CSOs in a single interview, the Horowitz survey differs from the Bortz methodology in that executives were queried about *all* distant signals transmitted by each of the cable systems. Based on the data provided by Horowitz, the number of distant signals associated with any one cable system ranged from one to over fifty; respondents would have been queried about all of the distant signals transmitted by their respective cable system. In contrast, Bortz

limited the number of distant signals for which a respondent had to report to eight (Bortz, p. 33-36).

49. As a result of their data collection approach, the Horowitz data are populated by a relatively small number of respondents. Table 1 shows the number of CSOs, the number of respondents, and the concentration of CSO responses for the Horowitz data. See also Trautman Written Rebuttal Testimony, Table A-4.

Table 1. Number of CSOs, Respondents, and Measures of Respondent Concentration, by Year, Horowitz Data

Year	Number of CSOs for which there are sample data	Number of respondents providing data for the CSOs in Column 2	Number of respondents reporting for 10+ CSOs	Percentage of data accounted for by the respondents in Column 4	Percent of data accounted for by the top 2 respondents
2010	123	31	3	42.4%	32.6%
2011	182	43	4	37.8%	25.2%
2012	228	42	8	58.9%	26.8%
2013	200	41	7	62.0%	29.0%

Source: Trautman Written Rebuttal Testimony, Table A-4.

As is evident from the table, especially for 2012 and 2013, a small number of individuals account for a large percentage of the data. And, as is evident from the final column of Table 1, in each year, two respondents account for more than a quarter of the data. The concentration of data exhibited in Table 1 is detrimental for two reasons: (1) the observations in the data are clearly not independent and should not be treated as such in the calculation of means and standard errors; and (2) with only two respondents accounting for over 25% of the data each year, these individuals can have an undue influence on the final estimates.

50. According to the methodology described by Horowitz (p. 8), when cable systems offered the same mix of distant signals, executives were to be interviewed once

concerning all of the similar CSOs. However, when I examine the data for a single respondent in a given year, I also find identical program valuations across CSOs with *varying* distant signals. For example, in 2013, looking only at the *sample* data used in estimation, respondent 54 (as identified in the Program Suppliers' data) provided information on 38 different cable systems.[13] For 15 of these 38 cable systems, the program valuations were as follows:

- News: 0% valuation
- Syndicated Series: 30%
- Movies: 15%
- Live Sports: 5%
- Other Sports: 0%
- Devotional: 0%
- Public Television (PTV): 50%
- Canadian: No valuation

However, the distant signals carried by these 15 cable systems varies, with no two cable systems offering the same mix of distant signals. It is quite surprising that this executive produced the *exact same valuations* for each of these 15 cable systems carrying different line-ups—assuming that he or she was interviewed separately about each system. Nor is this an isolated example; I see the same pattern of identical valuations for executives required to report for multiple cable systems across all four years of data.[14] These repeated identical responses regarding systems with non-identical signal lineups raise questions as to whether the survey protocol for separate questionnaires was in fact

[13] JSC_2010_2013_Masked_withDistantStations_MSOchanges_13July2017.xlsx.

[14] The example provided above is with respect to the repetition seen among those cases identified as part of the Horowitz sample (used for estimation by Dr. Frankel). The pattern of identical reporting across cable systems is even more evident when one looks at the full universe of systems for which a single executive was asked to report.

correctly implemented—or whether some respondents employed "short-cuts" in response to the burden of being asked to respond for numerous systems.

E. Summary of Horowitz Survey's Design Problems

51. The survey as designed and implemented by Dr. Horowitz and which forms the basis of the estimates provided by Dr. Frankel is fraught with problems. These problems include, but are not limited to:

- The extensive use of misleading and incorrect examples in the program category descriptions as well as the inconsistent use of the "such as" program examples;
- The failure to provide information identifying compensable programs on WGNA;
- The addition of an inappropriate "other sports programming" category;
- The consolidation of surveys in which a respondent was queried about multiple systems simultaneously; and
- The implementation of a data collection methodology that was excessively burdensome in that it requested respondents to report not only on sampled cable systems but all cable systems as well as reporting for *all* distant signals associated with each of the cable systems.

The extent of the misinformation provided as examples or as subcategories of programs ("such as") in the program category descriptions and the inconsistent use of examples and subcategories raises serious questions as to the validity of the responses and resulting estimates of program category valuations. Diamond (2011) notes that "[w]hen unclear questions are included in a survey, they may threaten the validity of the survey by systematically distorting responses if respondents are misled in a particular direction, or by inflating random error if respondents guess because they do not understand the

question" (p. 388). In this case, I believe that the provision of misinformation (exacerbated by the failure to provide information related to compensable programming) is sufficiently egregious as to reject the estimates of relative valuations resulting from the Horowitz survey. As a result of the issues I have outlined above, the Horowitz data provide neither a valid or reliable basis on which to estimate program valuations.

F. Data Adjustments

52. For those cable systems for which PBS was the only distant signal, the Horowitz questionnaire asks the following: "Considering the value of the programs broadcast only on PBS station (INSERT PBS STATIONS) to your cable system, what percentage, if any, of the fixed dollar amount would you allocate for this type of programming" (Horowitz, Appendix A, p. 36). PBS-only cable system executives were not instructed that the value of their estimate needed to add to 100%.

53. The question, as posed, is confusing, because how is an executive to value a program category relative to other categories if the cable system only offers programming in a single category, in this case, PBS? Regardless, the questionnaire does allow respondents to provide answers less than 100%. Such answers are clearly evident in the Horowitz survey responses. There are several[15] cases for which PTV-only systems reported valuations less than 100% for the PTV category. For example, in 2012, the relative program valuations for the 20 PTV-only systems range from 2% to 75%. However, it appears that Dr. Frankel adjusted these values to equal 100% (see, for

[15] See JSC_2010_2013_Masked_withDistantStations_MSOchanges_13July2017.xlsx. In 2010, 3 of the 15 (20%) of the PTV-only cable systems had valuations less than 100%; in 2011, 28 of the 28 (100%) of the PTV-only cable system had valuations less than 100%; in 2012, 20 of the 20 (100%) PTV-only cable systems had valuations less than 100%; and in 2013, 20 of the 20 (100%) of the PTV-only stations had valuations less than 100%.

example, the "reproportion" line of code in MPAA_2012.f90). Dr. Frankel provides no justification for altering the reported valuation.

G. Comparison of Statistical Estimates

54. The CRJs have in prior distribution orders cited the importance of focusing on confidence intervals around an estimate as opposed to strict adherence to the point estimates (Federal Register, Vol. 75, September 17, 2010, pp. 57066, 57068). Table IV-2 of the Bortz report provides 95% confidence intervals for the seven program categories used in the Bortz survey.

55. Dr. Frankel in his written direct testimony provides standard errors for the estimates derived from the Horowitz survey, rather than 95% confidence intervals. In order to provide an apples-to-apples comparison of the two sets of estimates, I have set forth below the point estimates, the margin of error[16], and the 95% confidence intervals for the Horowitz-based surveys, along with the 95% confidence intervals produced in Table IV-2 of the Bortz report.

[16] Margin of error = standard error of the estimate x 1.96, where 1.96 is the value corresponding to an alpha level of .05, that is, a 95% confidence level.

Table 2. Point Estimates, Margin of Error and 95% Confidence Intervals for Distant Signal Programming Valuation, by Programming Type, Survey Organization, and Year (95% confidence interval in parentheses)

	Bortz	**Horowitz**
2010		
Live professional and college team sports	40.9% ±1.6% (39.3% - 42.5%)	31.9 ±4.25 (27.7% - 36.2%)
Other sports	N/A	6.8% ± 1.3% (5.5% - 8.0%)
News and public affairs	18.7% ± 1.2% (17.5% - 19.9%)	12.4% ±2.9% (9.5% - 15.3%)
Movies	15.9% ± 0.7% (15.2%-16.6%)	17.2% ± 2.3% (14.9% - 19.4%)
Syndicated shows, series and specials	16.0% ± 1.0% (15.0%-16.9%)	20.3% ± 3.3% (16.9% - 23.6%)
PBS and all other programming on non-commercial signals	4.4% ± 0.9% (3.6%-5.3%)	7.7% ± 3.3% (4.4% - 11.0%)
Devotional and religious programming	4.0% ± 0.4% (3.6% ±4.4%)	3.8% ±1.5% (2.3% - 5.3%)
All programming on Canadian signals	0.1% ± 0.1% (0.0% - 0.2%)	0.0% ± 0.0% (0.0% - 0.0%)
2011		
Live professional and college team sports	36.4% ± 1.4% (34.9% - 37.8%)	27.1% ± 3.0% (24.1% - 30.1%)
Other sports	N/A	10.8% ± 1.6% (9.3% - 12.3%)
News and public affairs	18.3% ± 1.2% (17.1% - 19.6%)	12.9% ± 2.0% (10.9% - 14.8%)
Movies	18.6% ± 0.9% (17.7% - 19.5%)	11.4% ± 1.6% (9.9% - 13.0%)
Syndicated shows, series and specials	17.4% ± 1.0%) (16.3% - 18.4%)	17.6% ± 2.1% (15.5% - 19.7%)
PBS and all other programming on non-commercial signals	4.7% ±0.9% (3.9% - 5.6%)	13.3% ± 3.3% (10.1% - 16.6%)
Devotional and religious programming	4.5% ± 0.4% (4.1% - 4.9%)	5.9% ± 1.3% (4.6% - 7.2%)
All programming on Canadian signals	0.2% ± 0.1% (0.0% - 0.3%)	1.0% ± 1.7% (0.0% - 2.7%)

2012		
Live professional and college team sports	37.9% ± 1.8% (36.1% - 39.7%)	25.5% ± 2.9% (22.6% - 28.4%)
Other sports	N/A	9.0% ± 1.3% (7.7% - 10.3%)
News and public affairs	22.8% ± 1.0% (21.8% - 23.8%)	15.7% ± 1.7% (14.0% - 17.4%)
Movies	15.3% ±0.8% (14.5% - 16.1%)	12.1% ± 1.4% (10.7% - 13.6%)
Syndicated shows, series and specials	13.5% ± 0.6% (12.9% - 14.1%)	16.0% ± 2.0% (14.0% - 18.0%)
PBS and all other programming on non-commercial signals	5.1% ± 0.8% (4.3% - 5.9%)	15.1% ± 3.6% 11.5% - 18.6%
Devotional and religious programming	4.8% ± 0.4% (4.4% - 5.2%)	5.7% ± 0.8% (5.0% - 6.5%)
All programming on Canadian signals	0.6% ± 0.6% (0.1% - 1.2%)	0.9% ± 0.7% (0.2% - 1.6%)
2013		
Live professional and college team sports	37.7% ± 1.2% (36.4% - 38.9%)	35.3% ±9.5% (25.8% - 44.8%)
Other sports	N/A	7.4% ± 1.5% (5.9% - 8.9%)
News and public affairs	22.7% ± 1.0% (21.7% - 23.6%)	9.5% ± 2.0% (7.6% - 11.5%)
Movies	15.5% ± 0.8% (14.7% - 16.2%)	12.4% ± 2.5% (9.9% - 14.9%)
Syndicated shows, series and specials	11.8% ± 0.7% (11.0% - 12.5%)	16.3% ± 3.1% (13.1% - 19.4%)
PBS and all other programming on non-commercial signals	6.2% ± 0.8% (5.4% - 7.0%)	15.4% ± 6.6% (8.8% - 22.0%)
Devotional and religious programming	5.1% ± 0.3% (4.8% - 5.4%)	3.5% ± 0.9% (2.6% - 4.3%)
All programming on Canadian signals	1.2% ± 0.9% (0.4% - 2.1%)	0.4% ± 0.3% (0.1% - 0.6%)

Note: Data sources for Table 2 include Direct Testimony of Martin R. Frankel, Tables 5-8 (pp. 8 and 9) for the Horowitz column and Tables IV-1 (p. 42), IV-2 (p. 44), and Appendix D (pp. D-8 through D-11) for the Bortz column. Computation of margin of error and the 95% confidence interval for the Horowitz data computed by N. Mathiowetz based on the standard errors presented by Dr. Frankel. All estimates rounded to one significant digit. In 2010, the Horowitz estimate for all programming on Canadian Signals was 0.01% which rounds to 0.0% as presented in this table.

56. Looking only at the data for 2013 (for illustrative purposes), we see significant differences in the valuations for news and public affairs, syndicated shows, series, and specials, PTV, and devotional programming. The wider confidence intervals seen in the Horowitz-produced data renders several of the comparisons non-significant. For example, looking at live professional and college team sports for 2013, the 95% confidence interval produced from the Bortz data is 36.4% to 38.9%—a spread of ± 1.2 percentage points—whereas the interval produced from the Horowitz data is 25.8% to 44.8%—a spread of ± 9.5 percentage points.

V. THE FORD/RINGOLD SURVEY DOES NOT PROVIDE A RELIABLE BASIS FOR ALLOCATING RELATIVE VALUE TO CANADIAN PROGRAMMING

57. The written direct testimony of Debra J. Ringold summarizes the methodology and estimates resulting from the Ford/Ringold survey of U.S. cable system operators who retransmitted Canadian television stations as distant signals in 2010 through 2013. The Ford/Ringold survey design is similar to that used by Bortz and Horowitz in which a sample of cable system operators are interviewed about the relative value the operator would assign to categories of programs using a constant sum methodology. However, there are significant differences with respect to the sample design and the precision of the estimates between the Ford/Ringold survey and the Bortz survey.

58. The Ford/Ringold design indicates that CSOs were interviewed about "one Canadian signal randomly chosen from those Canadian signals retransmitted" (CCG-6, p. 4) but no information is provided as to how the signal was selected. It appears that the sample design of for the Ford/Ringold survey gave preference to French-language signals ("If cable systems were found to retransmit both an English-language and French-

language Canadian signal, the system was interviewed with the French-language version of the questionnaire, due to the smaller number of French-language systems" CCG-6, p. 6). As a result of this preference, the resulting analytic sample over-represents French-language systems. Whereas French-language stations accounted for about 21% of distant subscriber instances in 2013 (see CCG-1, Table 1 and Table 2, pp. 2-3, 5), the composition of the Ford/Ringold analytic sample consists of between 36% to 55% French-Language systems (computation based on data provided in CCG-6, Table 5 and CCG-6, Table 6).[17]

59. Diamond (2011) asks, "Does the sample approximate the relevant characteristics of the population?" In the case of the Ford/Ringold sample design, the analytic sample clearly over-represents a segment of the population, that is the French-language stations.

60. Diamond (2011) also notes that "all sample surveys produce estimates of population values, not exact measures of those values" (p. 381). One factor that affects the margin of error around a survey estimate is the size of the analytic sample. In the case of the study completed by Drs. Ford and Ringold, the sample sizes are extremely small, leading to large 95% confidence intervals for those estimates. Listed below are the estimates for the average value of the programming reported by Drs. Ford and Ringold in Table 1 (CCG-6, p. 15) for the "live professional and college team sports" category. Table 3 includes my computation of the standard errors as well as the 95% confidence interval of the estimates, under the assumption of a simple random sample.

[17] Specifically for 2010-2013, the proportion of French-Language Canadian Signals in the Ford/Ringold analytic sample is 38% (8 of 21), 44% (8 of 18), 36% (5 of 14) and 55% (6 of 11), respectively.

Table 3. Average Value of Live Professional and College Team Sports Shown on Canadian Signals with Standard Errors and 95% Confidence Intervals

Year	Estimate produced by Drs. Ford and Ringold (Table 1) (Sample size in parentheses)	Standard Deviation produced by Drs. Ford and Ringold (Table 1)	Standard Error of the Estimate	95% Confidence Interval (based on the standard error of the estimate)
2010	26.67 (21)	18.05	3.94	18.45 to 34.88
2011	14.72** (18)	9.92**	2.35**	10.14 to 19.30**
2012	21.07 (14)	21.23	5.67	8.81 to 33.33
2013	20.91 (11)	17.72	5.34	9.01 to 32.83

** My analysis of the Ford/Ringold data indicates that for 2011, the average value of live professional and college team sports is 15.52 with a standard deviation of 10.26, a standard error of 2.34 and a 95% confidence interval of 10.58 to 20.47

61. Two points of interest. First, Drs. Ford and Ringold produced standard deviations of the estimates, not standard errors. A standard deviation measures the dispersion of a set of data whereas a standard error is a measure of the reliability of an estimate. The two measures are related in that the standard error of an estimate is equal to the standard deviation of the estimate divided by the square root of the sample size. The 95% confidence interval, as described by Diamond (2011) "means that if 100 samples of the same size were drawn, the confidence interval expected for at least 95 of the samples would be expected to include the true population value" (p. 381). It does not mean that one is 95% confident that the true population value falls within the range provided. Second, in contrast to the Bortz survey, we see that the small sample size for the

Ford/Ringold survey leads to highly unreliable estimates (that is, wide confidence intervals).

62. The over-representation of French-speaking channels, coupled with the unreliable estimates, render the data from the Ford/Ringold study to be of little to no utility with respect to the issue of relative market value of Canadian programming on Canadian distant signals retransmitted by cable system operators in the United States.

63. Beyond the problems outlined above, a secondary issue with respect to the report of Drs. Ford and Ringold is the production of importance estimates for programming on TBS, U.S. superstations, and U.S. independent stations. Drs. Ford and Ringold note that the assessment of the relative importance of programming on these stations was conducted "to reduce the chances that respondents would guess the survey purpose or sponsor" (CCG-6, p. 4). Although I am supportive of the goal of masking the survey's purpose and sponsorship to respondents, the introduction of program categories that differ from those related to the primary purpose of the study adds unnecessarily to the cognitive burden of the respondents. Rather than simply reporting on the one constant sum question of interest before the CRJs, respondents to the Ford/Ringold survey were queried with respect to (up to) three different sets of program categories. This additional burden was unnecessary and may have led to confusion on the part of the respondents when reporting on the key question of interest, the relative programming value for Signal B stations.

64. Grouping together superstations such as WGN and WPIX with the cable network TBS likely led to additional confusion. Apart from the fact that TBS is not a distant signal, several of the program categories included in the constant sum question for

Signal A cable systems are irrelevant to TBS (news, children's programming, and devotional categories). Asking respondents to report on the relative value of programming that is not even offered would most likely further confuse respondents. According to the data produced by Drs. Ford and Ringold, of the 42 times that respondents were queried about a "superstation," 68.9% of the respondents were answering the questions with respect to TBS.

65. Similar to the estimates for the Canadian distant signals, the estimates for superstations (Table 2, CCG-6, p. 16) and independent stations (Table 3, CCG-6, p. 17) are based on very small sample sizes and are therefore subject to wide confidence intervals (unreliable estimates). Table 4 provides the standard errors and 95% confidence intervals for the live professional and college team sports based on the means and standard deviations produced by Drs. Ford and Ringold.

Table 4. Average Value of Live Professional and College Team Sports Shown on "superstations" and independent stations with Standard Errors and 95% Confidence Intervals

Year	Estimate produced by Drs. Ford and Ringold (Table 2 or 3) (Sample size in parentheses)	Standard Deviation produced by Drs. Ford and Ringold (Table 2 or 3)	Standard Error of the Estimate	95% Confidence Interval (based on the standard error of the estimate)
Superstation Estimates				
2010	35.00 (19)	20.75	4.76	25.67 to 44.33
2011	26.76 (17)	11.58	2.81	21.26 to 32.26
2012	19.64 (14)	12.32	3.29	13.19 to 26.09
2013	23.50 (10)	16.17	5.11	13.48 to 33.52
Independent Estimates				
2010	16.25 (4)	17.97	8.99	-1.37 to 33.87
2011	25.00 (5)	16.58	7.41	10.47 to 39.53
2012	24.00 (5)	4.18	1.87	20.33 to 27.66
2013	31.67 (3)	14.43	8.33	15.34 to 48.00

Note: Standard errors and confidence intervals produced for comparison purposes only. It is my usual practice to not produce estimates or confidence intervals when the number of observations within a cell is below n=20.

Similar to the estimates of Canadian distant signals, the unreliability of the estimates renders them uninformative with respect to understanding program valuations for superstations and independent stations.

I declare under penalty of perjury that the foregoing is true and correct.

Executed on __ September 14, 2017.

Nancy A. Mathiowetz

Nancy A. Mathiowetz, Ph.D.

Proof of Delivery

I hereby certify that on Thursday, October 17, 2019, I provided a true and correct copy of the ancy Mathiowetz Written Rebuttal Testimony (JSC Written RebuttalTestimony Vol. II) to the following:

Settling Devotional Claimants, represented by Matthew J MacLean, served via Electronic Service at matthew.maclean@pillsburylaw.com

Major League Soccer, LLC, represented by Edward S. Hammerman, served via Electronic Service at ted@copyrightroyalties.com

Multigroup Claimants, represented by Brian D Boydston, served via Electronic Service at brianb@ix.netcom.com

MPA-represented Program Suppliers, represented by Theresa B. Bowman, served via Email

Broadcaster Claimants Group (CTV), represented by David J Ervin, served via Electronic Service at dervin@crowell.com

Spanish Language Producers, represented by Brian D Boydston, served via Electronic Service at brianb@ix.netcom.com

Signed: /s/ Michael E Kientzle